Simultaneous Pairwise Multiple Comparisons in a Two-Way Design with Fixed Concomitant Variables

by

Ying-wang WONG

A Thesis Submitted to the

Graduate School

of

The Chinese University of Hong Kong

(Division of Statistics)

In Partial Fulfillment

of the Requirements for the Degree of

Master of Philosophy

(M. Phil.)

HONG KONG

December, 1996

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undersigned certify that we have read a thesis, entitled "Simultaneous Pairwise Multiple Comparisons in a Two-Way Design with Fixed Concomitant Variables" submitted to the Graduate School by Ying-wang Wong (黄 英 宏) in partial fulfillment of the requirement for the degree of Master of Philosophy in Statistics. We recommend that it be accepted.

Cheung S. H.

Dr. S. H. Cheung,
Supervisor

Wu Ka Ho

Dr. K. H. Wu

Leung Pui Lam

Dr. P. L. Leung

Dr. Paul S. F. Yip
External Examiner

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

ACKNOWLEDGEMENT

I am deeply indebted to my supervisor, Dr. S. H. Cheung, for his strict supervision and kind assistance during the course this research program. Without his valuable advice, this thesis would not be come to alive. I am grateful to Professors Paul S. F. Yip, K. H. Wu and P. L. Leung for their patience in reading the final draft and providing several helpful comments which led to a significant improvement of this thesis.

I would like to express my sincere gratitude to all the staff of the Department of Statistics for their generous assistance.

Finally, I want to thank my family for their support and encouragement, and especially to my wife, Ida, for her earnest support.

Y. W. WONG

Department of Statistics
The Chinese University of Hong Kong
Hong Kong
December, 1996

ABSTRACT

Tukey's (1953, The Problem of Multiple Comparisons, unpublished report, Princeton University) procedure is widely used for pairwise multiple comparisons in one-way Analysis of Variance (ANOVA). Copenhaver and Holland (1987, 1988) extended Tukey's procedure to balanced two-way designs. Cheung and Chan (1996) generalized the above procedure to the unbalanced case. In this thesis, pairwise multiple comparison procedures are developed for two-way Analysis of Covariance (ANCOVA) designs. The discussion will be limited to fixed covariates. In experimental design, the major advantage of using ANCOVA techniques is to reduce the error term variance. Numerical examples will be presented to illustrate how the proposed procedure is being applied to obtain exact simultaneous pairwise confidence intervals.

KEY WORDS AND PHRASES: Fixed Covariates, Simultaneous Pairwise Comparisons, Multivariate Normal Density Function

Contents

1. Introduction

1.1 Multiple Comparisons Procedures

Multiple testing is a widely used data analysis technique in scientific experiments. As indicated by Westfall and Young (1993), there are three major reasons for its popularity. The first reason relates to the cost of obtaining additional information. In most circumstances, it is more economical to collect information on various aspects of the same experimental units than to perform several experiments separately. The second reason is to detect interesting relationships not suspected *a priori*. 'Data splitting' which examines the data set in many different ways is extremely popular, especially with the help of modern computing technology and statistical softwares. Finally, researchers are always confronted with different kinds of compatible statistical techniques which also provide a source of multiple testing.

The most alarming problem of multiple testing procedures is that users may not be aware of the multiplicity effect. When a large data set is under extensive data splitting without a careful control of the overall error rate, 'false significance' can easily be obtained (Tukey, 1977; Diaconis, 1985). In the case of multiple hypotheses testing, the overall Type I error rate (at least one false rejection among all the hypotheses being considered) can be substantially large as explained in the following example.

Consider a random sample $\{X_1, X_2, \cdots, X_n\}$ with $X_i \sim N(\mu, \sigma^2)$, $i = 1, \cdots, n$. The $100(1-\alpha)\%$ confidence interval for μ is

$$(\overline{X} - \frac{S}{\sqrt{n}} t_{n-1,\alpha/2}, \overline{X} + \frac{S}{\sqrt{n}} t_{n-1,\alpha/2}) \tag{1.1.1}$$

where $\overline{X}$ and S^2 are the sample mean and sample variance respectively and $t_{n-1,\alpha/2}$ is the upper $\alpha/2$ percentage point of the t-distribution with $n-1$ degrees of freedom. Now assume that we have two independent random samples $\{X_1, X_2, \cdots, X_n\}$ and $\{Y_1, Y_2, \cdots, Y_m\}$. Furthermore, let $X_i \sim N(\mu_1, \sigma_1^2)$, $i = 1, \cdots, n$; $Y_j \sim N(\mu_2, \sigma_2^2)$, $j = 1, \cdots, m$. By (1.1.1), the $100(1-\alpha)\%$ confidence interval for μ_1 is $(\overline{X} - \frac{S_1}{\sqrt{n}} t_{n-1,\alpha/2}, \overline{X} + \frac{S_1}{\sqrt{n}} t_{n-1,\alpha/2})$ and the $100(1-\alpha)\%$ confidence interval for μ_2 is $(\overline{Y} - \frac{S_2}{\sqrt{m}} t_{m-1,\alpha/2}, \overline{Y} + \frac{S_2}{\sqrt{m}} t_{m-1,\alpha/2})$. However, if two such samples are obtained, for example, the examination scores of n boys and m girls from the same class, the research interest may be to estimate μ_1 and μ_2 simultaneously. In such cases what will be the joint coverage probability which provides a probability that both confidence intervals cover the true parameters? The joint coverage probability in this case drops to $(1-\alpha)^2$. The coverage probability is different if the two samples are dependent. If a large number of simultaneous confidence intervals are constructed and each has a nominal confidence level, say 0.95, the joint coverage probability will be extremely small. Similarly, if multiple hypotheses testing is performed, the overall error probability will be very much inflated if the researcher only put the focus on the control of error probability of each hypothesis.

To tackle the multiplicity problem, multiple comparison procedures (MCP) which control the overall error rate should be employed. MCP can be applied to eliminate the erroneous conclusions and reduce the cost of design in fulfilling the same experimental objectives. But the use of MCP depends on how we define the family or group of hypotheses in a particular problem. There are different types of families. For example, a collection of all pairwise comparisons in a two-way analysis of variance can be considered as a *finite* family. A collection of any linear contrast in an one-way analysis of variance forms an *infinite* family. In fact, the definition of family depends on the researcher and the type of experiment the researcher is doing. Cox (1965) suggests the following two important criteria for regarding a set of inferences as a family:

(a) To take into the account the multiple effect due to data-snooping.

(b) To ensure that the simultaneous correctness of a set of inferences so as to guarantee a correct overall decision.

Miller (1981) concludes that *"There are no hard-and-fast rules for where the family lines should be drawn, and the statistician must rely on own judgement for the problem at hand"*.

1.2 Familywise Error Rate

Tukey (1953) and Hochberg and Tamhane (1987) suggested that MCP are designed to take into account and properly control for multiplicity effect through some combined or joint measures of erroneous inferences. Hochberg and Tamhane (1987) considered three types of error rates. Per-comparison error rate (PCE) is defined as the expected proportion of incorrect inferences. Familywise error rate (FWE) is the probability of making any error in the given family of inferences. Per-family error rate (PFE) is the expected number of error in the family. They further explained the concept of a family as a collection of inferences for which it is meaningful to take into account some combined measure of errors.

To control the error rate in a given multiple comparison problem, Tukey (1953) preferred controlling the FWE to PFE with the following reasons:

(a) Control of the FWE for the family of all potential inferences ensures that the probability of any error in the selected set of inferences is controlled.

(b) For an infinite family the FWE can be controlled but not the PFE.

(c) When the requirement of the simultaneous correctness of all inferences must be satisfied, the FWE is the only choice for control.

Hochberg and Tamhane (1987) suggested that when stricter measure are needed in hypothesis testing, the FWE appeared to be more appropriate because control of FWE at level α provides an upper bound of α on the family of hypotheses. By controlling the FWE on the studies, researchers do not need to report the selection process in details. As a result, the proposed procedures in this thesis are constructed to control the FWE of Type I errors.

1.3 One-step Procedures Versus Stepwise Procedures

MCP can be classified into two groups: one-step procedures and stepwise procedures. One-step procedures are procedures that use the same critical value for all comparisons in a family of hypotheses without a predetermined order of testing. Stepwise procedures, on the other hand, employ a sequence of steps, each depending on the ones before it. An example of the one-step procedures is the Dunnett's (1955) procedure for comparisons of all active treatments with a control while maintaining a designated overall Type I error rate α. A well-known example of the stepwise procedures is the Fisher's protected least significance difference test (Fisher (1935)). There are two advantages of the one-step procedures. First, it is usually easier to implement. Second, through inversion, one can obtain the confidence intervals in a straightforward fashion. For detail comparison of these two types of procedures, one can see Hochberg and Tamhane (1987).

1.4 Pairwise Multiple Comparisons

Assumed in an experiment, we have c treatments. A comparison of c means is a linear combination of the means, such as the difference between two of the c means, or the difference between one mean and the average of two other means. Multiple comparisons of the c means can be expressed as

$$\Phi = l_1\mu_1 + l_2\mu_2 + \cdots + l_c\mu_c = \sum_{j=1}^{c} l_j\mu_j \tag{1.4.1}$$

where $\sum_{j=1}^{c} l_j = 0$ with at least one $l_j \neq 0$.

The most common one-step MCP are pairwise comparisons, all linear combination comparisons, comparisons with a control, comparisons with the best and comparisons with the average.

Scheffé's S-procedure (Scheffé (1953)) is a famous representation of all linear combination comparisons. It accepts any combinations of the l_j's in (1.4.1). Pairwise comparisons are in the form of (1.4.1) by placing weights of 1 and -1 to two of the r means and zero for all others. The most widely used procedure which controls the FWE is the Tukey-Kramer procedure (Tukey, 1953; Kramer, 1956; Kramer,1957). Multiple comparisons with a control is the case where one of the c treatments is designated as a control and different treatment groups are compared with it. A popular technique is the Dunnett (1955) procedure.

Among various procedures, pairwise comparisons and multiple comparisons with the control are definitely the most frequently applied methods. However, in this thesis, the focus will be limited to pairwise multiple comparisons.

1.5 Pairwise Multiple Comparisons in Two-Way Designs

Given a design with two qualitative factors (A and B), consider the linear fixed effect model

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \qquad i = 1, \ldots, r;\ j = 1, \ldots, c;\ k = 1, \ldots, n_{ij}; \tag{1.5.1}$$

with $\varepsilon_{ijk} \overset{ind}{\sim} N(0, \sigma^2)$. Let μ_{ij} denote the mean of the j^{th} level of factor B at the i^{th} level of factor A. The model (1.5.1) can be rewritten as

$$y_{ijk} = \mu + \tau_i + \gamma_j + \tau\gamma_{(ij)} + \varepsilon_{ijk} \tag{1.5.2}$$

where $\mu_{ij} = \mu + \tau_i + \gamma_j + \tau\gamma_{(ij)}$ with constraints $\sum_{i=1}^{r} \tau_i = 0$, $\sum_{j=1}^{c} \gamma_j = 0$, $\sum_{i=1}^{r} \tau\gamma_{(ij)} = 0$ and $\sum_{j=1}^{c} \tau\gamma_{(ij)} = 0$. The parameter μ is the overall mean effect, whereas τ_i is the i^{th} treatment effect of factor A, γ_j is the j^{th} treatment effect of factor B, $\tau\gamma_{(ij)}$ is the corresponding interaction effect.

It is common to perform pairwise multiple comparisons on one of the factors (for instance, factor B). For illustration, let us consider the following example. We have six different new drugs (factor B) to treat a certain disease. An experiment was conducted to examine their effects on both male and female (factor A) patients. In the presence of interaction between drugs and sex (different drugs behave very differently for the two genders), pairwise comparisons of the drugs based on the treatment means averaged across different levels of factor A (gender) is inappropriate. A more reasonable way is to perform the comparisons separately for these two groups (male and female). From the perspective of a researcher who is equally interested in treatment efficacy for both genders, he should be more conservative and control the familywise Type I error rate over all thirty comparisons.

To fix notation with model (1.5.1), pairwise comparisons refer to the simultaneous inferences of

$$\mu_{ij_1} - \mu_{ij_2} \tag{1.5.3}$$

for all $i = 1, \ldots, r$ and $1 \leq j_1 \neq j_2 \leq c$. Therefore, there are altogether $rc(c-1)/2$ comparisons. Methodology was developed by Copenhaver and Holland (1987, 1988) for the balanced case where all n_{ij} ($i = 1, \ldots, r$; $j = 1, \ldots, c$) are equal. Hence model (1.5.1) is simplified to

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \qquad i = 1, \ldots, r;\ j = 1, \ldots, c;\ k = 1, \ldots, n; \tag{1.5.4}$$

To construct procedures of simultaneous inferences of (1.5.3), the pivotal statistics

$$\frac{\sqrt{n}(\bar{y}_{ij_1} - \bar{y}_{ij_2}) - (\mu_{ij_1} - \mu_{ij_2})}{\hat{\sigma}} \tag{1.5.5}$$

($i = 1, \ldots, r$; $1 \leq j_1 \neq j_2 \leq c$) is employed and $\bar{y}_{ij} = \sum_{k=1}^{n} y_{ijk}$. The estimator $\hat{\sigma}^2$ is an unbiased estimator of σ^2. For a given α, the simultaneous $100(1-\alpha)\%$ confidence intervals for (1.5.3) are

$$(\bar{y}_{ij_1} - \bar{y}_{ij_2}) \pm Q'_{(\alpha,r,c,\nu)}\hat{\sigma}\sqrt{\frac{1}{n}} \tag{1.5.6}$$

($i = 1, \ldots, r$; $1 \leq j_1 \neq j_2 \leq c$) where $Q'_{(\alpha,r,c,\nu)}$ is the value of Q' which satisfies the following equation

$$P\left(\frac{\sqrt{n}(\bar{y}_{ij_1} - \bar{y}_{ij_2}) - (\mu_{ij_1} - \mu_{ij_2})}{\hat{\sigma}} \leq Q', \quad i = 1, \ldots, r; 1 \leq j_1 \neq j_2 \leq c\right) = 1 - \alpha. \tag{1.5.7}$$

Note that when $r = 1$, $Q'_{(\alpha,r,c,\nu)}$ reduces to the upper α point of the Studentized Range Distribution with parameter c and degrees of freedom ν. Selected values of $Q'_{(\alpha,r,c,\nu)}$ are tabulated in Copenhaver and Holland (1988) and practical examples can be found in their papers.

For the unbalanced case, the statistical procedures were given in Cheung and Chan (1996). Statistical procedures related to multiple comparisons with a control in a two-way design can be found in Cheung and Holland (1991, 1992, 1994). When $r = 1$, all the above mentioned procedure will be reduced to the one-way multiple comparison techniques.

1.6 Objectives

The major objective of this thesis is to generalize the pairwise multiple comparison procedures discussed in Section 1.5 to the case where concomitant variables (covariates) are included in the model. A review of the pairwise comparison procedures in one-way designs with covariates and the advantages of using covariates will be given in Section 2. In Section 3, pairwise multiple comparison methods with fixed covariates in two-way designs will be examined. Our proposed procedures can be classified as an one-step procedure as explained in Section 1.3. The extension to stepwise procedures will be left for further researches which are beyond the scope of this thesis. Finally in Section 4, numerical examples will be employed to illustrate our proposed procedures.

2 Pairwise Multiple Comparisons in One-Way Design with Covariates

2.1 The General ANCOVA Model

Consider the analysis of covariance (ANCOVA) model

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\mu} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.1.1}$$

where $\mathbf{y}$ is the $N \times 1$ response vector and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2)$. The $N \times p$ matrix $\mathbf{Z}$ is the design matrix corresponding to the $p \times 1$ vector $\boldsymbol{\mu}' = (\mu_1, \ldots, \mu_p)'$. Hence, each column of $\mathbf{Z}$ corresponds to a treatment. The $N \times q$ matrix $\mathbf{X}$ is the regression matrix corresponding to the $q \times 1$ vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)'$. Therefore, $\mathbf{Z}\boldsymbol{\mu}$ represents the qualitative part and $\mathbf{X}\boldsymbol{\beta}$ represents the quantitative (covariates) part of the right hand side of model (2.1.1). We will make the assumptions that both $\mathbf{Z}$ and $\mathbf{X}$ have full column rank and the columns of $\mathbf{Z}$ are linearly independent of those of $\mathbf{X}$.

For model (2.1.1), an example will be an one-way design with 3 treatments, each having 5 observations and one covariate. Then, we have

$$\begin{aligned}
\mathbf{y} &= (y_{11}, y_{12}, y_{13}, y_{14}, y_{15}, y_{21}, \cdots, y_{25}, y_{31}, \cdots, y_{35})'_{15\times 1} \\
\boldsymbol{\mu} &= (\mu_1, \mu_2, \mu_3)'_{3\times 1} \\
\boldsymbol{\beta} &= (\beta)_{1\times 1} \\
\boldsymbol{\varepsilon} &= (\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{13}, \varepsilon_{14}, \varepsilon_{15}, \varepsilon_{21}, \cdots, \varepsilon_{25}, \varepsilon_{31}, \cdots, \varepsilon_{35})'_{15\times 1}
\end{aligned}$$

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix}_{15\times 3}, \qquad \mathbf{X} = \begin{pmatrix} x_{11} - \bar{x}_{..} \\ x_{12} - \bar{x}_{..} \\ x_{13} - \bar{x}_{..} \\ x_{14} - \bar{x}_{..} \\ x_{15} - \bar{x}_{..} \\ x_{21} - \bar{x}_{..} \\ \vdots \\ x_{25} - \bar{x}_{..} \\ x_{31} - \bar{x}_{..} \\ \vdots \\ x_{35} - \bar{x}_{..} \end{pmatrix}_{15\times 1}.$$

The normal equation of (2.1.1) is

$$\begin{pmatrix} \mathbf{Z'Z} & \mathbf{Z'X} \\ \mathbf{X'Z} & \mathbf{X'X} \end{pmatrix} \begin{pmatrix} \mu \\ \beta \end{pmatrix} = \begin{pmatrix} \mathbf{Z'Y} \\ \mathbf{X'Y} \end{pmatrix} \tag{2.1.2}$$

where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\beta}}$ are the least squares estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$ respectively. Solving (2.1.2), we have

$$\hat{\beta} = (\mathbf{X'PX})^{-1}\mathbf{X'Py} \tag{2.1.3}$$

and

$$\hat{\mu} = (\mathbf{Z'Z})^{-1}\mathbf{Z}'(\mathbf{y} - \mathbf{X}\hat{\beta}) \tag{2.1.4}$$

where

$$\mathbf{P} = \mathbf{I} - \mathbf{Z}(\mathbf{Z'Z})^{-1}\mathbf{Z}'. \tag{2.1.5}$$

Let the variance of $\hat{\boldsymbol{\beta}}$ be $\boldsymbol{\Lambda}\sigma^2$ and it can be shown that

$$\boldsymbol{\Lambda} = (\mathbf{X'PX})^{-1}. \tag{2.1.6}$$

The variance of $\hat{\boldsymbol{\mu}}$ is

$$Var\,(\hat{\boldsymbol{\mu}}) = \boldsymbol{\Sigma}\sigma^2 \tag{2.1.7}$$

where

$$\begin{aligned} \boldsymbol{\Sigma} &= (\mathbf{Z'Z})^{-1} + (\mathbf{Z'Z})^{-1}\mathbf{Z'X}(\mathbf{X'PX})^{-1}\mathbf{X'Z}(\mathbf{Z'Z})^{-1} \\ &= (\mathbf{Z'Z})^{-1} + (\mathbf{Z'Z})^{-1}\mathbf{Z'X}\boldsymbol{\Lambda}\mathbf{X'Z}(\mathbf{Z'Z})^{-1}. \end{aligned} \tag{2.1.8}$$

Hence, with the assumption that $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2)$, $\hat{\boldsymbol{\mu}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}\sigma^2)$. Furthermore, the estimate of σ^2 is

$$\hat{\sigma}^2 = \frac{\mathbf{y'P}\left[\mathbf{I} - \mathbf{X}\,(\mathbf{X'PX})^{-1}\,\mathbf{X'P}\right]\mathbf{y}}{\nu} \tag{2.1.9}$$

with $\nu = N - r(\mathbf{Z}) - r(\mathbf{X})$ where $r(\mathbf{Z})$ and $r(\mathbf{X})$ are the column ranks of $\mathbf{Z}$ and $\mathbf{X}$ respectively. The estimator $\hat{\sigma}^2$ is distributed independently of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\beta}}$ and

$$\frac{\nu\hat{\sigma}^2}{\sigma^2} \sim \chi^2$$

with ν degrees of freedom. One may refer to Searle (1971) for details.

A close inspection on the model (2.1.1) reveals an interesting property that the model is composed of a quantitative part (covariates) and a qualitative part (treatments). Therefore, ANCOVA model can be thought of a combination of a regression model and an analysis of variance model.

In experimental designs, the major advantage of using ANCOVA is its ability to reduce the error term variance (Scheffé (1959), Neter *et. al.* (1990)). In the presence of strong linear relationship between the dependent variable and the covariates, the observations can be well adjusted by the covariates. This reduces the variability of the random error. With the use of covariates, the relationship between the treatments stands out more clearly. However, one should be careful in choosing the covariates. If the treatments do affect the covariates, this relationship may be blurred or even wrong.

The covariates ($\mathbf{X}$) in the above model can be either fixed or random. When the covariates are fixed, they are predetermined or being controlled in the experiment. For example, consider the study of an enzyme activity in different types of substrate under temperature of $0°$, $20°$, $40°$, $60°$. In this case the covariate, temperature, is fixed since it is predetermined before the experiment.

There are cases where the covariates cannot be controlled. For example, a study of the sales ($\mathbf{Y}$) of the ice-cream of different brands on the mean daily temperature ($\mathbf{X}$) in a supermarket. As the temperature of the day cannot be controlled, it is more suitable to consider both $\mathbf{Y}$ and $\mathbf{X}$ as random variables.

In this thesis, we will limit ourselves to fixed covariates and let us proceed to review the statistical methodologies in multiple comparisons with the presence of fixed covariates in one-way designs.

2.2 Pairwise Comparisons

In this section, we will introduce methods which are related to fixed covariates. For procedures which handle random covariates, one may refer to Thigpen and Paulson (1974), Bryant and Paulson (1976), Bryant and Bruvold (1980) and Hochberg and Varon-Salomon (1984).

Several popular procedures are discussed here briefly. They are the Fisher protected LSD (least significant difference) procedure, the Bonferroni procedure, the Scheffé procedure and the Tukey-Kramer approximation method.

These procedures can be applied to balanced and unbalanced designs. Among these procedures, Fisher protected LSD procedure is restricted to hypothesis testing only, while the other three procedures can be employed in hypothesis testing and construction of simultaneous confidence intervals.

We first examine the Fisher protected LSD procedure (Fisher, 1935). Let the null hypothesis be $\mu_{j_1} - \mu_{j_2} = 0$, $1 \leq j_1 \neq j_2 \leq c$. Assume $\mathbf{l}$ be the corresponding contrast vector such that $\mathbf{l}'\boldsymbol{\mu} = \mu_{j_1} - \mu_{j_2}$, then the test statistics is

$$\frac{\mathbf{l}'\hat{\boldsymbol{\mu}}}{\hat{\sigma}\sqrt{\mathbf{l}'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{l} + \boldsymbol{\theta}'\boldsymbol{\Lambda}\boldsymbol{\theta}}} = t_\nu \qquad (2.2.1)$$

where $\boldsymbol{\theta}' = \mathbf{l}'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}$ and t_ν has Student's t distribution with ν degrees of freedom. The pairwise comparison tests are performed if the initial ANCOVA F test is significant. It is a stepwise procedure and simultaneous confidence intervals are not available.

The Bonferroni procedure is based on the popular Bonferroni first order inequality. It is powerful if we have a small number of planned pairwise or contrast comparisons. The $100(1-\alpha)\%$ confidence intervals for the set of planned comparisons $\{\mathbf{l}_1'\boldsymbol{\mu}, \mathbf{l}_2'\boldsymbol{\mu}, \cdots, \mathbf{l}_m'\boldsymbol{\mu}\}$ are

$$\mathbf{l}_i'\hat{\boldsymbol{\mu}} \pm t_{(\frac{\alpha}{2m},\nu)}\hat{\sigma}\sqrt{\mathbf{l}_i'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{l}_i + \boldsymbol{\theta}_i'\boldsymbol{\Lambda}\boldsymbol{\theta}_i} \qquad (1 \leq i \leq m) \tag{2.2.2}$$

where $t_{(\frac{\alpha}{2m},\nu)}$ is the upper percentage point of the t distribution with ν degrees of freedom and $\boldsymbol{\theta}_i' = \mathbf{l}_i'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}$.

The Scheffé procedure (Scheffé, 1953), on the other hand, is good for a family of infinite number of comparisons. The $100(1-\alpha)\%$ simultaneous confidence intervals for $\mathbf{l}'\boldsymbol{\mu}$ ($\forall\,\mathbf{l}' \neq \mathbf{0}$) are

$$\mathbf{l}'\hat{\boldsymbol{\mu}} \pm \sqrt{(c-1)F_{(\alpha,c-1,\nu)}\hat{\sigma}^2\left(\mathbf{l}'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{l} + \boldsymbol{\theta}'\boldsymbol{\Lambda}\boldsymbol{\theta}\right)}. \tag{2.2.3}$$

Since the Scheffé method allows a set of infinite number of comparisons with a predetermined confidence level, it is hence expected to yield relatively conservative intervals, especially when the actual number of intervals required is small.

By the Inversion Theorem (see for example p. 407, Casella and Berger, 1990), both intervals (2.2.2) and (2.2.3) can be used to implement multiple hypothesis testing with the predetermined overall Type I error rate α.

To construct pairwise simultaneous confidence intervals (or simultaneous testing procedures), the extension of the Tukey-Kramer (Tukey, 1953; Kramer, 1956, 1957) approximation method provides an alternative to existing procedures. The $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_{j_1} - \mu_{j_2}$, $1 \leq j_1 \neq j_2 \leq c$ are

$$\hat{\mu}_{j_1} - \hat{\mu}_{j_2} \pm Q_{(\alpha,c,\nu)}\frac{\hat{\sigma}}{\sqrt{2}}\sqrt{\frac{1}{n_{j_1}} + \frac{1}{n_{j_2}} + \boldsymbol{\theta}_{(j_1,j_2)}'\boldsymbol{\Lambda}\boldsymbol{\theta}_{(j_1,j_2)}} \tag{2.2.4}$$

for $1 \leq j_1 \neq j_2 \leq c$ and $\boldsymbol{\theta}_{(j_1,j_2)}' = \mathbf{l}_{(j_1,j_2)}'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}$ where $\mathbf{l}_{(j_1,j_2)}'\boldsymbol{\mu} = \mu_{j_1} - \mu_{j_2}$. The quantity $Q_{(\alpha,c,\nu)}$ is the upper α point of the Studentized Range Distribution with parameters c and degrees of freedom ν. Note that the conservative property of the above procedure is not yet being established (see for example Hayter, 1989).

3. Pairwise Comparisons in Two-Way Layout With Covariates

3.1 The Model

Similar to model (1.5.1), the model with two qualitative factors and q covariates can be written as

$$y_{ijk} = \mu_{ij} + \beta_1(x_{1ijk} - \bar{x}_{1\ldots}) + \cdots + \beta_q(x_{qijk} - \bar{x}_{q\ldots}) + \varepsilon_{ijk}, \tag{3.1.1}$$

where $i = 1, \ldots, r$; $j = 1, \ldots, c$; $k = 1, \ldots n_{ij}$, with $\varepsilon_{ijk} \overset{ind}{\sim} N(0, \sigma^2)$. Furthermore, x_{wijk} represents the value of the w^{th} covariate ($w = 1, \ldots, q$) corresponding to y_{ijk}. Let $\bar{x}_{w\ldots} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n_{ij}} \frac{x_{wijk}}{N}$ where $N = \sum_{i=1}^{r}\sum_{j=1}^{c} n_{ij}$ is the total sample size.

Let $\boldsymbol{\mu}_i' = (\mu_{i1}, \mu_{i2}, \cdots, \mu_{ic})$ for $i = 1, \ldots, r$. Corresponding to model (2.1.1),

$$\boldsymbol{\mu}' = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_2', \cdots, \boldsymbol{\mu}_r')$$

$$\boldsymbol{\beta}' = (\beta_1, \beta_2, \cdots, \beta_q)$$

and $\underset{N\times rc}{\mathbf{Z}}$ is a diagonal matrix with diagonal elements ($\mathbf{J}_{11}$, $\mathbf{J}_{12}, \cdots$, $\mathbf{J}_{1c}$, $\mathbf{J}_{21}$, $\mathbf{J}_{22}, \cdots$, $\mathbf{J}_{2c}, \cdots$, $\mathbf{J}_{r1}$, $\mathbf{J}_{r2}, \cdots$, $\mathbf{J}_{rc}$) where $\mathbf{J}_{ij}$ denotes a column vector of n_{ij} ones ($i = 1, \ldots, r$; $j = 1, \ldots, c$). Hence, $\mathbf{Z}$ is a full column rank matrix. Let the estimates of $\boldsymbol{\mu}'$ and $\boldsymbol{\beta}'$ be $\hat{\boldsymbol{\mu}}' = (\hat{\boldsymbol{\mu}}_1', \hat{\boldsymbol{\mu}}_2', \cdots, \hat{\boldsymbol{\mu}}_r')$ and $\hat{\boldsymbol{\beta}}' = (\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_q)$ respectively. It is straightforward to show that

$$\hat{\mu}_{ij} = \bar{y}_{ij.} - \sum_{w=1}^{q} \hat{\beta}_w \left(\bar{x}_{wij.} - \bar{x}_{w\ldots}\right)$$

with $\bar{y}_{ij.} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk}$ and $\bar{x}_{wij.} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} x_{wijk}$. Since $\mathrm{Cov}\left[\left(\mathbf{Z}'\mathbf{Z}\right)^{-1} \mathbf{Z}'\mathbf{y}, \hat{\boldsymbol{\beta}}\right] = \mathbf{0}$, $\bar{y}_{ij.}$ is independent of $\hat{\beta}_w$ ($1 \le i \le r$, $1 \le j \le c$, $1 \le w \le q$). The pairwise differences $\mu_{ij_1} - \mu_{ij_2}$, $1 \le i \le r$, $1 \le j_1 \le j_2 \le c$ can be estimated by the unbiased estimator $\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}$ and

$$Var\left(\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}\right) = \left(\frac{1}{n_{ij_1}} + \frac{1}{n_{ij_2}} + \boldsymbol{\theta}_{i(j_1,j_2)}' \boldsymbol{\Lambda} \boldsymbol{\theta}_{i(j_1,j_2)}\right) \sigma^2 \tag{3.1.2}$$

where

$$\boldsymbol{\theta}_{i(j_1,j_2)}=\begin{pmatrix} \bar{X}_{1ij_1.} - \bar{X}_{1ij_2.} \\ \bar{X}_{2ij_1.} - \bar{X}_{2ij_2.} \\ \vdots \\ \bar{X}_{qij_1.} - \bar{X}_{qij_2.} \end{pmatrix}.$$

3.2 The Test Statistics

In order to perform simultaneous pairwise comparisons of $\mu_{ij_1} - \mu_{ij_2}$ for all $1 \le i \le r$, $1 \le j_1 < j_2 \le c$, we define the pivotal statistics

$$T_{i(j_1,j_2)} = \frac{(\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}) - (\mu_{ij_1} - \mu_{ij_2})}{\hat{\sigma}\sqrt{d_{i(j_1,j_2)}}} \tag{3.2.1}$$

where

$$d_{i(j_1,j_2)} = \frac{1}{n_{ij_1}} + \frac{1}{n_{ij_2}} + \boldsymbol{\theta}'_{i(j_1,j_2)}\boldsymbol{\Lambda}\boldsymbol{\theta}_{i(j_1,j_2)}. \tag{3.2.2}$$

To conduct hypothesis testing of the $rc(c-1)/2$ null hypotheses

$$H_0: \ \mu_{ij_1} = \mu_{ij_2} \tag{3.2.3}$$

versus the two sided alternatives

$$H_1: \ \mu_{ij_1} \neq \mu_{ij_2}$$

for all $1 \le i \le r$, $1 \le j_1 < j_2 \le c$ with familywise Type I error rate α or to construct the two-sided $100(1-\alpha)\%$ simultaneous pairwise confidence intervals for mean differences $\mu_{ij_1} - \mu_{ij_2}$, we need to find the value of t such that

$$P\left\{\left|T_{i(j_1,j_2)}\right| \le t; \qquad 1 \le i \le r,\ 1 \le j_1 < j_2 \le c\right\} = 1 - \alpha. \tag{3.2.4}$$

The solution of t for equation (3.2.4) will be denoted by t_α.

With familywise Type I error rate α, each hypothesis in (3.2.3) is rejected if and only if the test statistic $\left|T_{i(j_1,j_2)}\right|$ ($\mu_{ij_1} - \mu_{ij_2} = 0$ under the null hypothesis) exceeds t_α; equivalently, if

$$|\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}| > t_\alpha \hat{\sigma}\sqrt{d_{i(j_1,j_2)}}. \tag{3.2.5}$$

The corresponding two-sided $100(1-\alpha)\%$ confidence intervals for mean differences $\mu_{ij_1} - \mu_{ij_2}$ are

$$(\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}) \pm t_\alpha \hat{\sigma}\sqrt{d_{i(j_1,j_2)}} \tag{3.2.6}$$

for all $1 \leq i \leq r$, $1 \leq j_1 < j_2 \leq c$.

3.3 Computation of Upper Percentage Points

Now, let us proceed to solve equation (3.2.4) for the value of t. Note that

$$\begin{aligned} & P\left\{\left|T_{i(j_1,j_2)}\right| \leq t; \qquad 1 \leq i \leq r,\ 1 \leq j_1 < j_2 \leq c\right\} \\ = \;& P\left\{\left|\frac{(\hat{\mu}_{ij_1}-\hat{\mu}_{ij_2})-(\mu_{ij_1}-\mu_{ij_2})}{\hat{\sigma}\sqrt{d_{i(j_1,j_2)}}}\right| \leq t; \quad 1 \leq i \leq r,\ 1 \leq j_1 < j_2 \leq c\right\} \end{aligned} \tag{3.3.1}$$

Let $u = \dfrac{\hat{\sigma}}{\sigma}$ and $g(u)$ be the density function of $\sqrt{\dfrac{\chi^2_\nu}{\nu}}$. Conditioning on $\dfrac{\hat{\sigma}}{\sigma}$, (3.3.1) becomes

$$\int_0^\infty \left[P\left\{\left|\frac{(\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}) - (\mu_{ij_1} - \mu_{ij_2})}{\sigma\sqrt{d_{i(j_1,j_2)}}}\right| \leq ut;\ 1 \leq i \leq r,\ 1 \leq j_1 < j_2 \leq c\right\}\right] g(u)du \tag{3.3.2}$$

Since $\hat{\mu}_{ij_1} = \bar{y}_{ij_1.} - \sum_{w=1}^{q} \hat{\beta}_w(\bar{X}_{wij_1.} - \bar{X}_{w...})$, $\hat{\mu}_{ij_2} = \bar{y}_{ij_2.} - \sum_{w=1}^{q} \hat{\beta}_w(\bar{X}_{wij_2.} - \bar{X}_{w...})$. Hence

$$\begin{aligned} & \hat{\mu}_{ij_1} - \hat{\mu}_{ij_2} \\ = \;& \left[\left(\bar{y}_{ij_1.} - \sum_{w=1}^{q} \beta_w\left(\bar{X}_{wij_1.} - \bar{X}_{w...}\right)\right) - \left(\bar{y}_{ij_2.} - \sum_{w=1}^{q} \beta_w\left(\bar{X}_{wij_2.} - \bar{X}_{w...}\right)\right)\right] \end{aligned}$$

$$- \left[\left(\sum_{w=1}^{q} \hat{\beta}_w \left(\bar{X}_{wij_1.} - \bar{X}_{w...} \right) - \sum_{w=1}^{q} \beta_w \left(\bar{X}_{wij_1.} - \bar{X}_{w...} \right) \right) \right.$$

$$\left. - \left(\sum_{w=1}^{q} \hat{\beta}_w \left(\bar{X}_{wij_2.} - \bar{X}_{w...} \right) - \sum_{w=1}^{q} \beta_w \left(\bar{X}_{wij_2.} - \bar{X}_{w...} \right) \right) \right]$$

$$= \quad (\bar{y}_{ij_1.} - \bar{y}_{ij_2.}) - \sum_{w=1}^{q} \beta_w \left(\bar{X}_{wij_1.} - \bar{X}_{wij_2.} \right) - \sum_{w=1}^{q} \left(\hat{\beta}_w - \beta_w \right) \left(\bar{X}_{wij_1.} - \bar{X}_{wij_2.} \right)$$

Note that

$$\frac{1}{\sigma} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) = \frac{1}{\sigma} \begin{pmatrix} \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_2 - \beta_2 \\ \vdots \\ \hat{\beta}_q - \beta_q \end{pmatrix} \sim N(\mathbf{0}, \boldsymbol{\Lambda})$$

where $\boldsymbol{\Lambda} = \{\lambda_{ij}\}, 1 \le i, j \le q$. Writing explicitly, $\boldsymbol{\Lambda}^{-1}$ is a $q \times q$ matrix with elements $\sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n_{ij}} (x_{q_1' ijk} - \bar{x}_{q_1' ij.})(x_{q_2' ijk} - \bar{x}_{q_2' ij.})$, and $1 \le q_1', q_2' \le q$. Let

$$\boldsymbol{\Gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_q \end{pmatrix} = \frac{1}{\sigma} \begin{pmatrix} \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\lambda_{11}}} \\ \frac{\hat{\beta}_2 - \beta_2}{\sqrt{\lambda_{22}}} \\ \vdots \\ \frac{\hat{\beta}_q - \beta_q}{\sqrt{\lambda_{qq}}} \end{pmatrix}$$

and $h(\gamma_1, \gamma_2, \cdots, \gamma_q)$ be the joint density of $(\gamma_1, \gamma_2, \cdots, \gamma_q)$. Hence $(\gamma_1, \gamma_2, \cdots, \gamma_q)$ have a multivariate normal distribution with mean vector $\mathbf{0}$ and variance covariance matrix $\mathbf{R}$ which is the correlation matrix of $\frac{1}{\sigma} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right)$. Conditioning on $\boldsymbol{\Gamma}$, (3.3.2) becomes

$$\int_0^{\infty} \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P \left\{ \left| \frac{(\bar{y}_{ij_1.} - \bar{y}_{ij_2.}) - \sum_{w=1}^{q} \beta_w \left(\bar{X}_{wij_1.} - \bar{X}_{wij_2.} \right) - (\mu_{ij_1} - \mu_{ij_2})}{\sigma \sqrt{d_{i(j_1, j_2)}}} \right. \right. \right.$$

$$- \sum_{w=1}^{q} \gamma_w \sqrt{\lambda_{ww}} \frac{\left(\bar{X}_{wij_1.} - \bar{X}_{wij_2.}\right)}{\sqrt{d_{i(j_1,j_2)}}} \Bigg| \le ut; \quad 1 \le i \le r,\ 1 \le j_1 < j_2 \le c \Bigg\}$$

$$h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \Bigg]\ g(u)\, du \tag{3.3.3}$$

Let

$$Z_{ij} = \frac{(\bar{y}_{ij.} - \bar{y}_{ic.}) - \sum\limits_{w=1}^{q} \beta_w \left(\bar{X}_{wij.} - \bar{X}_{wic.}\right) - (\mu_{ij} - \mu_{ic})}{\sigma\sqrt{d_{i(j,c)}}}, \quad 1 \le j \le c-1$$

and

$$D_{ij} = \sum_{w=1}^{q} \gamma_w \sqrt{\lambda_{ww}} \frac{\left(\bar{X}_{wij.} - \bar{X}_{wic.}\right)}{\sqrt{d_{i(j,c)}}}, \quad 1 \le j \le c-1,$$

the above probability inequality becomes

$$\int_0^\infty \Bigg[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P \Bigg\{ -ut + D_{ij_1} \le Z_{ij_1} \le ut + D_{ij_1}, \quad 1 \le j_1 \le c-1;$$

$$-ut + \frac{D_{ij_1}\sqrt{d_{i(j_1,c)}} - D_{ij_2}\sqrt{d_{i(j_2,c)}}}{\sqrt{d_{i(j_1,j_2)}}} \le \frac{Z_{ij_1}\sqrt{d_{i(j_1,c)}} - Z_{ij_2}\sqrt{d_{i(j_2,c)}}}{\sqrt{d_{i(j_1,j_2)}}} \le ut +$$

$$\frac{D_{ij_1}\sqrt{d_{i(j_1,c)}} - D_{ij_2}\sqrt{d_{i(j_2,c)}}}{\sqrt{d_{i(j_1,j_2)}}}; \quad 1 \le i \le r,\ 1 \le j_1 < j_2 \le c-1; \Bigg\}$$

$$h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \Bigg] g(u) du \tag{3.3.4}$$

By adopting the arguments of Hayter (1989) and Cheung and Chan (1996), (3.3.4) can be written as

$$\int_0^\infty \Bigg\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{r} \left[\int_{-ut+D_{i1}}^{ut+D_{i1}} \int_{L_{i2}}^{U_{i2}} \cdots \int_{L_{i(c-1)}}^{U_{i(c-1)}} f\left(z_{i1}, \ldots, z_{i(c-1)}\right) dz_{i1} \cdots dz_{i(c-1)} \right]$$

$$h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \Bigg\} g(u) du \tag{3.3.5}$$

where

$$
\begin{aligned}
U_{ij} &= U_{ij}(ut + D_{ij}, z_{i1}, \ldots, z_{i(j-1)}) \\
&= \min\left\{ ut + D_{ij}, ut\sqrt{\frac{d_{i(1,j)}}{d_{i(j,c)}}} + (z_{i1} - D_{i1})\sqrt{\frac{d_{i(1,c)}}{d_{i(j,c)}}} + D_{ij}, \ldots, \right. \\
&\qquad \left. ut\sqrt{\frac{d_{i(j-1,j)}}{d_{i(j,c)}}} + \left(z_{i(j-1)} - D_{i(j-1)}\right)\sqrt{\frac{d_{i(j-1,c)}}{d_{i(j,c)}}} + D_{ij} \right\}, \quad (3.3.6)
\end{aligned}
$$

and

$$
\begin{aligned}
L_{ij} &= L_{ij}(-ut + D_{ij}, z_{i1}, \ldots, z_{i(j-1)}) \\
&= \max\left\{ -ut + D_{ij}, -ut\sqrt{\frac{d_{i(1,j)}}{d_{i(j,c)}}} + (z_{i1} - D_{i1})\sqrt{\frac{d_{i(1,c)}}{d_{i(j,c)}}} + D_{ij}, \ldots, \right. \\
&\qquad \left. -ut\sqrt{\frac{d_{i(j-1,j)}}{d_{i(j,c)}}} + \left(z_{i(j-1)} - D_{i(j-1)}\right)\sqrt{\frac{d_{i(j-1,c)}}{d_{i(j,c)}}} + D_{ij} \right\} \quad (3.3.7)
\end{aligned}
$$

for $1 \le i \le r$, $2 \le j \le c-1$, and $f(z_{i1}, \ldots, z_{i(c-1)})$ denotes the joint density of the random variables $(z_1, \ldots, z_{(c-1)})$ which have a multivariate normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{V}_i = (v_{i(j_1,j_2)})$, $1 \le i \le r$, $1 \le j_1, j_2 \le c-1$, where

$$
v_{i(j_1,j_2)} = \begin{cases} \dfrac{\frac{1}{n_{ij_1}} + \frac{1}{n_{ic}}}{d_{i(j_1,c)}} & j_1 = j_2 \\[2ex] \dfrac{\frac{1}{n_{ic}}}{\sqrt{d_{i(j_1,c)} d_{i(j_2,c)}}} & j_1 \neq j_2. \end{cases}
$$

When α is given, the algorithm which computes (3.2.4) and hence solves for the value of t_α numerically, is outlined in Appendix A.

3.4 Approximation Procedure

As similar to the Tukey-Kramer approximation method discussed in Section (2.2), one can generalize the approximate Tukey-Kramer procedure to the two-way layout as follows.

When familywise Type I error rate α is given, each hypothesis in (3.2.3) is rejected if and only if the test statistic $\left|T_{i(j_1,j_2)}\right|$ ($\mu_{ij_1} - \mu_{ij_2} = 0$ under the null hypothesis) exceeds $Q'_{(\alpha,r,c,\nu)}$; equivalently, if

$$|\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}| > Q'_{(\alpha,r,c,\nu)}\frac{\hat{\sigma}}{\sqrt{2}}\sqrt{d_{i(j_1,j_2)}}. \tag{3.4.1}$$

The corresponding two-sided $100(1-\alpha)\%$ confidence intervals for mean differences $\mu_{ij_1} - \mu_{ij_2}$ are

$$(\hat{\mu}_{ij_1} - \hat{\mu}_{ij_2}) \pm Q'_{(\alpha,r,c,\nu)}\frac{\hat{\sigma}}{\sqrt{2}}\sqrt{d_{i(j_1,j_2)}} \tag{3.4.2}$$

for all $1 \le i \le r$, $1 \le j_1 < j_2 \le c$.

The value of $Q'_{(\alpha,r,c,\nu)}$ can be obtained from the tables of Copenhaver and Holland (1988). This approximation method should be satisfactory especially when the effect of the covariates are small which means that $\boldsymbol{\theta}'_{i(j_1,j_2)}\boldsymbol{\Lambda}\boldsymbol{\theta}_{i(j_1,j_2)}$ are small for $1 \le i \le r$; $1 \le j_1 < j_2 \le c$. In such circumstance, one can choose to avoid the complex numerical work which was outlined earlier.

3.5 Two-Way Layout with One Covariate

In many experimental situations, the number of covariates is usually small. So we consider the particular case where there is only one covariate ($q = 1$) as an illustration of our procedure. If $q = 1$, Model (3.1.1) reduces to

$$y_{ijk} = \mu_{ij} + \beta(x_{ijk} - \bar{x}_{...}) + \varepsilon_{ijk}, \quad i = 1, \ldots, r;\ j = 1, \ldots, c;\ k = 1, \ldots n_{ij}, \tag{3.5.1}$$

with $\varepsilon_{ijk} \overset{ind}{\sim} N(0, \sigma^2)$ and $\bar{x}_{...} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n_{ij}} x_{ijk}/N$. Since we have only one covariate in model (3.5.1), $\boldsymbol{\beta} = \beta$ is a 1×1 scalar. With $\mathbf{P}$ defined in (2.1.5), it is easy to show that

$$\mathbf{X}'\mathbf{P}\mathbf{X} = S_{xx} \tag{3.5.2}$$

and

$$\mathbf{X}'\mathbf{P}\mathbf{y} = S_{xy} \tag{3.5.3}$$

where

$$\begin{aligned}
S_{xx} &= \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n_{ij}}(x_{ijk} - \bar{x}_{ij.})^2 \\
S_{xy} &= \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n_{ij}}(x_{ijk} - \bar{x}_{ij.})(y_{ijk} - \bar{y}_{ij.}) \\
\bar{x}_{ij.} &= \sum_{k=1}^{n_{ij}} x_{ijk}/n_{ij} \\
\bar{y}_{ij.} &= \sum_{k=1}^{n_{ij}} y_{ijk}/n_{ij}
\end{aligned}$$

Hence, by equation (2.1.3),

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} \tag{3.5.4}$$

and by equation (2.1.4), the estimate of μ_{ij} is

$$\hat{\mu}_{ij} = \bar{y}_{ij.} - \hat{\beta}(\bar{x}_{ij.} - \bar{x}_{...}). \tag{3.5.5}$$

By equation (2.1.9), the estimate of σ^2 is

$$\hat{\sigma}^2 = \frac{S_{yy} - \frac{S_{xy}^2}{S_{xx}}}{\nu} \tag{3.5.6}$$

where $S_{yy} = \sum_i\sum_j\sum_k(y_{ijk} - \bar{y}_{ij.})^2$. Furthermore, from (3.2.2), we have

$$d_{i(j_1,j_2)} = \frac{1}{n_{ij_1}} + \frac{1}{n_{ij_2}} + \frac{(\bar{x}_{ij_1.} - \bar{x}_{ij_2.})^2}{S_{xx}}. \tag{3.5.7}$$

4. Numerical Examples

Example 1. One-Way Layout with Two Covariates

The first example is extracted from Huitema (1980, p.162) which is related to a behavioural objectives study. The response variable is the score of a biology achievement test (Y) and the treatments are three different types of study objectives on student achievement in freshmen biology. The two covariates are aptitude test scores (X_1) and academic motivation test scores (X_2). The data are shown in Table 4.1.

According to Section (2.1), we obtain the following statistics:

$$\begin{array}{llll}
\hat{\mu}_1 = 28.979 & \bar{y}_1 = 30.0 & \bar{x}_{11} = 52.0 & \bar{x}_{21} = 5.1 \\
\hat{\mu}_2 = 40.212 & \bar{y}_2 = 39.0 & \bar{x}_{12} = 47.0 & \bar{x}_{22} = 4.8 \\
\hat{\mu}_3 = 35.809 & \bar{y}_3 = 36.0 & \bar{x}_{13} = 49.0 & \bar{x}_{23} = 5.1 \\
\bar{x}_{1.} = 49.333 & \bar{x}_{2.} = 5.0 & \hat{\beta}_1 = 0.277 & \hat{\beta}_2 = 2.835
\end{array}$$

$$\hat{\sigma}^2 = 38.897 \qquad \nu = 25$$

$$\boldsymbol{\theta}_1' = (5.0 \quad 0.3) \quad \boldsymbol{\theta}_2' = (3.0 \quad 0.0) \quad \boldsymbol{\theta}_3' = (-2.0 \quad -0.3)$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} 2.974 \times 10^{-4} & -1.177 \times 10^{-3} \\ -1.177 \times 10^{-3} & 1.135 \times 10^{-2} \end{pmatrix}.$$

If the covariates are not included in the model, the formula for the estimate of σ^2 is different. Let the unbiased estimator of σ^2 be $\tilde{\sigma}^2$ if the covariates are excluded in the analysis. In such case,

$$\tilde{\sigma}^2 = \frac{\mathbf{y}'\mathbf{P}\mathbf{y}}{\nu'}$$

where $\nu' = N - r(\mathbf{Z})$. With the data in Example 1, $\tilde{\sigma}^2 = 130.963$. Hence, the reduction of error variance estimate is 70.3% and this indicates that the covariates are very useful in giving more precise estimates of treatment means because the confidence intervals will be a lot narrower.

To construct 95% simultaneous pairwise confidence intervals (SPCI), we employ four different methods. They are the Bonferroni procedure, the Scheffé procedure, the Tukey-Kramer approximation procedure and our proposed procedure in Section (3.2). The result is summarized in Table 4.2.

The table indicates that our procedure yields narrower confidence intervals than both the Bonferroni and the Scheffé procedures. The approximation method seems to be excellent in this case. As mentioned earlier, when the number of pairwise comparisons is small, the Bonferroni procedure performs better than the Scheffé procedure.

Table 4.1 Data for the behavourial objectives study extracted from Huitema (1980, p.162).

Types of Study Objectives								
I			II			III		
X_1	X_2	Y	X_1	X_2	Y	X_1	X_2	Y
29	3	15	22	3	20	33	2	14
49	3	19	24	2	34	45	1	20
48	2	21	49	4	28	35	5	30
35	5	27	46	4	35	39	4	32
53	5	35	52	5	42	36	3	34
47	9	39	43	4	44	48	8	42
46	3	23	64	8	46	63	8	40
74	7	38	61	7	47	57	4	38
72	6	33	55	6	40	56	9	54
67	8	50	54	5	54	78	7	56

Table 4.2 95% joint confidence intervals for pairwise comparisons of the adjusted mean scores of a biology achievement test on three different types of study objectives on student achievement in freshmen biology.

$\mu_{j_1} - \mu_{j_2}$	$\mu_1 - \mu_2$	$\mu_1 - \mu_3$	$\mu_2 - \mu_3$
Bonferroni (Eqn. (2.2.2))	$(-18.48, -3.99)$	$(-14.03, 0.37)$	$(-2.77, 11.57)$
Scheffé (Eqn. (2.2.3))	$(-18.58, -3.89)$	$(-14.14, 0.48)$	$(-2.87, 11.68)$
Tukey-Kramer (Eqn. (2.2.4))	$(-18.27, -4.20)$	$(-13.82, 0.16)$	$(-2.56, 11.37)$
Our procedure (Eqn. (3.2.6))	$(-18.27, -4.20)$	$(-13.82, 0.16)$	$(-2.56, 11.37)$

$\mu_j =$ adjusted population mean score of a biology achievement test on jth type of study objective.

$t_{(0.0083,25)} = 2.566.$

$F_{(0.05,2,25)} = 2.602.$

$\frac{Q_{(0.05,3,25)}}{\sqrt{2}} = 2.491.$

$t_{0.05} = 2.491.$

Example 2. Two-Way Layout with Two Covariates

This example is extracted from Steel and Torrie (1980, p. 326). It is a study of weight change of guinea pigs fed by forage grown in different soil treatments (fertilized or unfertilized in 4 different types of soil). There are two independent covariates which are initial weight (X_1) and forage consumed (X_2) of guinea pigs. The response variable is gain in weight (Y). The data are shown in Table 4.3.

According to Section (2.1), we obtain the following statistics:

$$
\begin{array}{llll}
\hat{\mu}_{11} = 266.793 & \hat{\mu}_{21} = 251.001 & \bar{y}_{11} = 264.333 & \bar{y}_{21} = 252.667 \\
\hat{\mu}_{12} = 142.847 & \hat{\mu}_{22} = 109.790 & \bar{y}_{12} = 141.333 & \bar{y}_{22} = 105.333 \\
\hat{\mu}_{13} = 201.003 & \hat{\mu}_{23} = 171.706 & \bar{y}_{13} = 236.667 & \bar{y}_{23} = 175.000 \\
\hat{\mu}_{14} = 226.337 & \hat{\mu}_{24} = 233.857 & \bar{y}_{14} = 203.667 & \bar{y}_{24} = 224.333
\end{array}
$$

$$
\begin{array}{llll}
\bar{x}_{111} = 242.667 & \bar{x}_{121} = 268.000 & \bar{x}_{211} = 1384.667 & \bar{x}_{221} = 1471.000 \\
\bar{x}_{112} = 266.333 & \bar{x}_{122} = 236.333 & \bar{x}_{212} = 1447.000 & \bar{x}_{222} = 1357.000 \\
\bar{x}_{113} = 245.667 & \bar{x}_{123} = 249.000 & \bar{x}_{213} = 1632.000 & \bar{x}_{223} = 1436.000 \\
\bar{x}_{114} = 272.333 & \bar{x}_{124} = 214.333 & \bar{x}_{214} = 1328.000 & \bar{x}_{224} = 1272.667
\end{array}
$$

$$
\bar{x}_{1..} = 249.333 \qquad \bar{x}_{2..} = 1416.042 \qquad \hat{\beta}_1 = -0.378 \qquad \hat{\beta}_2 = 0.159
$$

$$
\begin{array}{ll}
\boldsymbol{\theta}'_{1(1,2)} = (-23.667 \quad -62.333) & \boldsymbol{\theta}'_{2(1,2)} = (\ 31.667 \quad 114.000) \\
\boldsymbol{\theta}'_{1(1,3)} = (-3.000 \quad -247.333) & \boldsymbol{\theta}'_{2(1,3)} = (\ 19.000 \quad 35.000) \\
\boldsymbol{\theta}'_{1(1,4)} = (-29.667 \quad 56.667) & \boldsymbol{\theta}'_{2(1,4)} = (\ 53.667 \quad 198.333) \\
\boldsymbol{\theta}'_{1(2,3)} = (\ 20.667 \quad -185.000) & \boldsymbol{\theta}'_{2(2,3)} = (-12.667 \quad -79.000) \\
\boldsymbol{\theta}'_{1(2,4)} = (-6.000 \quad 119.000) & \boldsymbol{\theta}'_{2(2,4)} = (\ 22.000 \quad 84.333) \\
\boldsymbol{\theta}'_{1(3,4)} = (-26.666 \quad 304.000) & \boldsymbol{\theta}'_{2(3,4)} = (\ 34.667 \quad 163.333)
\end{array}
$$

$$
\hat{\sigma}^2 = 395.067 \qquad \nu = 14 \qquad \mathbf{\Lambda} = \begin{pmatrix} 5.294 \times 10^{-5} & -6.006 \times 10^{-6} \\ -6.006 \times 10^{-6} & 2.360 \times 10^{-6} \end{pmatrix}.
$$

Table 4.3 Data for the trial study extracted from Steel and Torrie (1980, p.326). (All measurements are in grams.)

Unfertilized Soil			Fertilized Soil		
X_1	X_2	Y	X_1	X_2	Y
		Miami silt loam			
220	1155	224	222	1326	237
246	1423	289	268	1559	265
262	1576	280	314	1528	256
		Plainfield fine sand			
198	1092	118	205	1154	82
266	1703	191	236	1250	117
335	1546	115	268	1667	117
		Almena silt loam			
213	1573	242	188	1381	184
236	1730	270	259	1363	129
288	1593	198	300	1564	212
		Carlisle peat			
256	1532	241	202	1375	239
278	1220	185	216	1170	207
283	1232	185	225	1273	227

Table 4.4 95% joint confidence intervals for pairwise comparisons of the adjusted mean gain in weight of the trial study of the guinea pigs on four different types of land and two different soil treatment.

$\mu_{ij_1} - \mu_{ij_2}$ $(1 \leq j_1 \neq j_2 \leq 4)$		Group i 1 (Unfertilized)	Group i 2 (Fertilized)
$\mu_{i1} - \mu_{i2}$	Our procedure	(70.43, 177.46)	(86.95, 195.47)
	Appr. procedure	(70.17, 177.73)	(86.68, 195.74)
$\mu_{i1} - \mu_{i3}$	Our procedure	(7.98, 123.60)	(26.06, 132.53)
	Appr. procedure	(7.69, 123.89)	(25.79, 132.80)
$\mu_{i1} - \mu_{i4}$	Our procedure	(−15.09, 96.00)	(−39.99, 74.28)
	Appr. procedure	(−15.37, 96.28)	(−40.28, 74.57)
$\mu_{i2} - \mu_{i3}$	Our procedure	(−116.44, 0.13)	(−115.04, −8.79)
	Appr. procedure	(−116.73, 0.42)	(−115.31, −8.53)
$\mu_{i2} - \mu_{i4}$	Our procedure	(−137.88, −29.10)	(−177.54, −70.59)
	Appr. procedure	(−138.15, −28.83)	(−177.81, −70.33)
$\mu_{i3} - \mu_{i4}$	Our procedure	(−90.49, 39.83)	(−117.10, −7.20)
	Appr. procedure	(−90.82, 40.15)	(−117.38, −6.93)

$\mu_{i1} =$ adjusted population mean gain in weight of the guinea pigs with forage grown on Miami silt loam in group i.

$\mu_{i2} =$ adjusted population mean gain in weight of the guinea pigs with forage grown on Plainfield fine sand in group i.

$\mu_{i3} =$ adjusted population mean gain in weight of the guinea pigs with forage grown on Almena silt loam in group i.

$\mu_{i4} =$ adjusted population mean gain in weight of the guinea pigs with forage grown on Carlisle peat in group i.

Our procedure employs equation (3.2.6) with $t_{0.05} = 3.246$.

Appr. procedure employs equation (3.4.2) with $\frac{Q'_{(0.05,2,4,14)}}{\sqrt{2}} = 3.263$.

With the data in Example 2, $\tilde{\sigma}^2 = 130.963$ and the reduction of error variance estimate is 61.0%. This indicates that the covariates are very useful and therefore should be included in the model.

Table 4.4 gives 95% simultaneous confidence intervals for the treatment adjusted mean differences. We can compare the two different methods, the Tukey-Kramer approximation procedure and our proposed procedure. From the table, we notice that our proposed procedure yields slightly narrower simultaneous confidence intervals. The differences between these two methods are not significant because $\boldsymbol{\theta}'_{i(j_1,j_2)}\boldsymbol{\Lambda}\boldsymbol{\theta}_{i(j_1,j_2)}$ are small for $1 \le i \le r$, $1 \le j_1 \ne j_2 \le c$.

Example 3. Two-Way Layout with Two Covariates (Artificial Data)

The third example is a copy of the second example except the data in the Example 2 are replaced by an artificial data set. The function of Example 3 is to demonstrate that in some circumstances, the approximation method is not close to the exact solution and hence our proposed procedure is recommended. In this example, the data of $\mathbf{X}$ are changed (Table 4.5) such that $\boldsymbol{\theta}'_{i(j_1,j_2)}\boldsymbol{\Lambda}\boldsymbol{\theta}_{i(j_1,j_2)}$ are relatively large compared to the previous example for some $1 \le i \le r$, $1 \le j_1 \ne j_2 \le c$. As indicated in equation (3.2.2), if $\boldsymbol{\theta}'_{i(j_1,j_2)}\boldsymbol{\Lambda}\boldsymbol{\theta}_{i(j_1,j_2)}$ are small for $1 \le i \le r$, $1 \le j_1 \ne j_2 \le c$, $d_{i(j_1,j_2)}$ will be approximately equal to $\frac{1}{n_{ij_1}} + \frac{1}{n_{ij_2}}$. In such cases, the approximation method will be close to our exact method. In this example, we notice that t_α is quite different from $\frac{Q'_{(\alpha,r,c,\nu)}}{\sqrt{2}}$ (Table 4.6).

According to Section (2.1), we obtain the following statistics:

$\hat{\mu}_{11} = 817.614$	$\hat{\mu}_{21} = 694.526$	$\bar{y}_{11} = 830.000$	$\bar{y}_{21} = 766.000$
$\hat{\mu}_{12} = 753.450$	$\hat{\mu}_{22} = 1021.236$	$\bar{y}_{12} = 879.667$	$\bar{y}_{22} = 876.333$
$\hat{\mu}_{13} = 815.997$	$\hat{\mu}_{23} = 866.806$	$\bar{y}_{13} = 837.667$	$\bar{y}_{23} = 862.333$
$\hat{\mu}_{14} = 813.829$	$\hat{\mu}_{24} = 928.541$	$\bar{y}_{14} = 866.667$	$\bar{y}_{24} = 793.333$

$\bar{x}_{111} = 246.000$	$\bar{x}_{121} = 261.333$	$\bar{x}_{211} = 1451.333$	$\bar{x}_{221} = 1537.667$
$\bar{x}_{112} = 269.667$	$\bar{x}_{122} = 209.667$	$\bar{x}_{212} = 1680.333$	$\bar{x}_{222} = 1173.667$
$\bar{x}_{113} = 239.000$	$\bar{x}_{123} = 247.000$	$\bar{x}_{213} = 1565.333$	$\bar{x}_{223} = 1369.333$
$\bar{x}_{114} = 282.333$	$\bar{x}_{124} = 212.667$	$\bar{x}_{214} = 1234.667$	$\bar{x}_{224} = 1182.667$

$\bar{x}_{1..} = 249.333$ $\quad \bar{x}_{2..} = 1416.042$ $\quad \hat{\beta}_1 = -0.378$ $\quad \hat{\beta}_2 = 0.159$

$$
\begin{aligned}
\boldsymbol{\theta}'_{1(1,2)} &= (-23.667 \quad -229.000) & \boldsymbol{\theta}'_{2(1,2)} &= (\ 51.667 \quad 364.000)\\
\boldsymbol{\theta}'_{1(1,3)} &= (\ 7.000 \quad -114.000) & \boldsymbol{\theta}'_{2(1,3)} &= (\ 14.333 \quad 168.333)\\
\boldsymbol{\theta}'_{1(1,4)} &= (-36.333 \quad 216.667) & \boldsymbol{\theta}'_{2(1,4)} &= (\ 48.667 \quad 355.000)\\
\boldsymbol{\theta}'_{1(2,3)} &= (\ 30.667 \quad 115.000) & \boldsymbol{\theta}'_{2(2,3)} &= (-37.333 \quad -195.667)\\
\boldsymbol{\theta}'_{1(2,4)} &= (-12.667 \quad 445.667) & \boldsymbol{\theta}'_{2(2,4)} &= (-3.000 \quad -9.000)\\
\boldsymbol{\theta}'_{1(3,4)} &= (-43.333 \quad 330.667) & \boldsymbol{\theta}'_{2(3,4)} &= (\ 34.333 \quad 186.667)
\end{aligned}
$$

$$
\hat{\sigma}^2 = 395.067 \qquad \nu = 14 \qquad \boldsymbol{\Lambda} = \begin{pmatrix} 1.865 \times 10^{-3} & -1.573 \times 10^{-4} \\ -1.573 \times 10^{-4} & 1.335 \times 10^{-4} \end{pmatrix}.
$$

In this example, $\tilde{\sigma}^2 = 575.50$ and the reduction of error variance estimate is 48.9%. This also indicates that the covariates are useful in giving more precise estimates of treatment means.

The confidence intervals with the approximation method are much longer than the confidence intervals with our proposed method and it demonstrates the superiority of our proposed procedure.

Table 4.5 Data for Example 3.

Unfertilized Soil			Fertilized Soil		
X_1	X_2	Y	X_1	X_2	Y
Miami silt loam					
240	1455	820	262	1526	784
246	1423	830	268	1559	793
252	1476	840	254	1528	721
Plainfield fine sand					
268	1692	865	205	1154	879
266	1703	884	216	1200	893
275	1646	890	208	1167	857
Almena silt loam					
243	1573	870	252	1381	884
236	1530	803	249	1363	871
238	1593	840	240	1364	832
Carlisle peat					
286	1252	894	202	1175	788
278	1220	849	216	1200	793
283	1232	857	220	1173	799

Table 4.6 95% joint confidence intervals for pairwise comparisons of the adjusted mean for Example 3.

$\mu_{ij_1} - \mu_{ij_2}$ $(1 \le j_1 \ne j_2 \le 4)$		Group i: 1 (Unfertilized)	2 (Fertilized)
$\mu_{i1} - \mu_{i2}$	Our procedure	(−69.16, 197.49)	(−536.91, −116.51)
	Appr. procedure	(−83.97, 212.30)	(−560.26, −93.16)
$\mu_{i1} - \mu_{i3}$	Our procedure	(−81.81, 85.05)	(−273.94, −70.62)
	Appr. procedure	(−91.08, 94.32)	(−285.23, −59.33)
$\mu_{i1} - \mu_{i4}$	Our procedure	(−169.75, 177.32)	(−438.44, −29.59)
	Appr. procedure	(−189.03, 196.60)	(−461.15, −6.88)
$\mu_{i2} - \mu_{i3}$	Our procedure	(−150.89, 25.80)	(30.25, 278.61)
	Appr. procedure	(−160.70, 35.61)	(16.45, 292.41)
$\mu_{i2} - \mu_{i4}$	Our procedure	(−332.81, 212.05)	(50.99, 134.40)
	Appr. procedure	(−363.08, 242.32)	(46.36, 139.03)
$\mu_{i3} - \mu_{i4}$	Our procedure	(−240.81, 245.14)	(−179.86, 56.39)
	Appr. procedure	(−267.80, 272.13)	(−192.98, 69.51)

μ_{i1} = adjusted population mean gain in weight of the guinea pigs with forage grown on Miami silt loam in group i.

μ_{i2} = adjusted population mean gain in weight of the guinea pigs with forage grown on Plainfield fine sand in group i.

μ_{i3} = adjusted population mean gain in weight of the guinea pigs with forage grown on Almena silt loam in group i.

μ_{i4} = adjusted population mean gain in weight of the guinea pigs with forage grown on Carlisle peat in group i.

Our procedure employs equation (3.2.6) with $t_{0.05} = 2.936$.

Appr. procedure employs equation (3.4.2) with $\frac{Q'_{(0.05,2,4,14)}}{\sqrt{2}} = 3.263$.

Appendix A - An Algorithm in Solving Equation (3.2.4) for the value of t_α

1) Input values of r, c, n_{ij} $(1 \le i \le r, 1 \le j \le c)$, α, x_{wijk} $(1 \le w \le q$, $(1 \le i \le r, 1 \le j \le c, 1 \le k \le n_{ij})$.

2) The secant method is used to solve equation (3.2.4) for t_α when α is given. the program terminates if the difference between successive iterates is less than err_1 (suggested value: 0.00001).

3) The outer integral related to the density function $g(u)$ is evaluated using subroutine QPROB of Copenhaver (1987) with 16-point Gauss-Legrendre composite quadrature. Let

$$E = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{r} \left[\int_{-ut+D_{i1}}^{ut+D_{i1}} \int_{L_{i2}}^{U_{i2}} \cdots \int_{L_{i(c-1)}}^{U_{i(c-1)}} f\left(z_{i1}, z_{i2}, \ldots, z_{i(c-1)}\right) dz_{i1}\, dz_{i2} \cdots dz_{i(c-1)} \right] h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q$$

Then (3.3.5) becomes

$$\int_0^\infty E\, g(u)\, du. \tag{A.1}$$

The above integral was divided into subintervals of length L and (A.1) is approximated by

$$\sum_{m=0}^{\infty} \int_{mL}^{mL+L} E\, g(u)\, du. \tag{A.2}$$

Then, the limits of integration were rescaled from $(mL, mL+L)$ to $(-1, 1)$ so that Gauss-Legrendre quadrature could be employed. The accuracy of this algorithm was compared with that of a 24-point quadrature using intervals of length $L/2$, with little difference in results in the sixth decimal places, *cf.* Copenhaver and Holland (1988).

4) Let

$$H_i = \int_{-ut+D_{i1}}^{ut+D_{i1}} \int_{L_{i2}}^{U_{i2}} \cdots \int_{L_{i(c-1)}}^{U_{i(c-1)}} f\left(z_{i1}, z_{i2}, \ldots, z_{i(c-1)}\right) dz_{i1}\, dz_{i2} \cdots dz_{i(c-1)}.$$

Therefore

$$E = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{i=1}^{r} H_i\right) h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q.$$

Then, H_i is evaluated by the numerical integration method by Genz (1992) which is outlined in Appendix B. The Monte Carlo algorithm of Genz (1992) can be employed with standard error for the estimate of H_i to be err_2 (suggested value: 0.001).

The value of E is estimated by E' where

$$|E' - E| \leq err_3$$

with

$$E' = \int_{-\eta}^{\eta} \int_{-\eta}^{\eta} \cdots \int_{-\eta}^{\eta} \left(\prod_{i=1}^{r} H_i\right) h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q$$

where $\infty > \eta \geq 0$ is a constant determined by the following equation

$$err_3 = 1 - \int_{-\eta}^{\eta} \int_{-\eta}^{\eta} \cdots \int_{-\eta}^{\eta} h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \tag{A.3}$$

Note that $0 \leq E \leq 1$ and $0 \leq \prod_{i=1}^{r} H_i \leq 1$ because both E and $\prod_{i=1}^{r} H_i$ are probabilities of some events. Hence

$$0 \leq E - E' \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{i=1}^{r} H_i\right) h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q$$

$$
\begin{aligned}
& - \int_{-\eta}^{\eta} \int_{-\eta}^{\eta} \cdots \int_{-\eta}^{\eta} \left(\prod_{i=1}^{r} H_i \right) h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \\
\leq \; & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \\
& - \int_{-\eta}^{\eta} \int_{-\eta}^{\eta} \cdots \int_{-\eta}^{\eta} h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \\
= \; & 1 - \int_{-\eta}^{\eta} \int_{-\eta}^{\eta} \cdots \int_{-\eta}^{\eta} h(\gamma_1, \gamma_2, \cdots, \gamma_q)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_q \\
= \; & err_3
\end{aligned}
$$

The value of err_3 is specified by the user (suggested value: 0.0001). The value of η is then computed by equation (A.3).

5) Finally, the value of E' can be obtained by the application of the Genz algorithm once again with standard error for the estimate of E', err_4 (suggested value: 0.001).

Appendix B - Evaluation of Multivariate Normal Probabilities

The heavy dependence of the statistical analysis on the multivariate normal distribution is that this probabilistic model approximates well the distributions of continuous measurements in many sampled populations. With this important reason, statistical theories related to this model are extensively developed. Moreover, exact mathematical treatment is usually obtainable through normal theory, the distribution of many statistics can be obtained exactly.

Recently, with the rapid development of computing machines, statistical methods with high dimensions are more feasible.

A random vector $\mathbf{X} = (X_1, X_2, \cdots, X_k)'$ is said to have a k-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ if its characteristic function $\Psi_{\mathbf{X}}(\mathbf{u}) = E\left(e^{i\mathbf{u}'\mathbf{x}}\right)$ is given by

$$\Psi_{\mathbf{X}}(\mathbf{u}) = \exp\left(i\mathbf{u}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u}\right) \tag{B.1}$$

where $i^2 = -1$. If $\boldsymbol{\Sigma}$ is positive definite, then the density function of $\mathbf{X}$ exists and is given by

$$f(\mathbf{x}) = (2\pi)^{-\frac{k}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}. \tag{B.2}$$

Let $\mathbf{x} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}$ be positive definite, $-\infty \le \mathbf{x} \le \infty$. To compute

$$\begin{aligned} F(\mathbf{a},\mathbf{b}) &= Pr(\mathbf{a} \le \mathbf{x} \le \mathbf{b}) \\ &= ((2\pi)^p|\boldsymbol{\Sigma}|)^{-1/2} \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_p}^{b_p} \exp(-\frac{1}{2}\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x})d\mathbf{x} \end{aligned} \tag{B.3}$$

where $\mathbf{a}' = (a_1, a_2, ..., a_p)$, $\mathbf{b}' = (b_1, b_2, ..., b_p)$, one can use the algorithm of Schervish (1984). However, this algorithm requires extremely long computation times for $p > 6$. Genz (1992) proposed a transformation technique which enables us to efficiently evaluate (B.3). His method is composed of a sequence of three transformations. Let $\mathbf{x} = \mathbf{C}\mathbf{y}$ where $\mathbf{y}' = (y_1, ..., y_p)$. The matrix $\mathbf{C}\mathbf{C}'$ is the Cholesky Decomposition of $\mathbf{\Sigma}$ and $\mathbf{C} = \{c_{ij}\}$ is a lower triangular matrix. Also, let $\mathbf{w}' = (w_1, ..., w_p)$. Then (B.3) can be transformed to

$$F(\mathbf{a}, \mathbf{b}) = (e_1 - d_1) \int_0^1 (e_2 - d_2) \cdots \int_0^1 (e_p - d_p) \int_0^1 d\mathbf{w}, \tag{B.4}$$

where

$$\begin{aligned}
d_1 &= \Phi(a_1/c_{11}), \qquad\qquad e_1 = \Phi(b_1/c_{11}) \\
w_i &= (\Phi(y_i) - d_i)/(e_i - d_i), \quad i = 1, ..., p \\
d_i &= \Phi((a_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(d_j + w_j(e_j - d_j)))/c_{ii}), \quad i = 2, ..., p \\
e_i &= \Phi((b_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(d_j + w_j(e_j - d_j)))/c_{ii}), \quad i = 2, ..., p
\end{aligned}$$

and Φ is the standard normal distribution function. After the above transformations, the evaluation of $F(\mathbf{a}, \mathbf{b})$ is far easier than with (B.3). If the mean of $\mathbf{x}$ is $\boldsymbol{\mu} \neq \mathbf{0}$, we can simply let $\mathbf{x} - \boldsymbol{\mu} = \mathbf{C}\mathbf{y}$ instead of $\mathbf{x} = \mathbf{C}\mathbf{y}$. Hence, the algorithm can be applied to the evaluation of multivariate normal probabilities with non-zero mean vectors. Genz (1992) reported that even a simple Monte Carlo algorithm is very effective, and other details of the algorithm can be found in Genz (1992).

The algorithm of Joe (1995) is an alternative to Genz (1992) algorithm. Both algorithms are efficient and highly accurate (For the comparisons of these two algorithms, see Joe (1995)).

REFERENCES

Bryant, J. L. and Bruvold, N. T. (1980), "Multiple Comparison Procedures in the Analysis of Covariance," *Journal of American Statistical Association*, **75**, 874-880.

Bryant, J. L. and Paulson, A. S. (1976), "An Extension of Tukey's Method of Multiple Comparisons to Experimental Designs with Random Concomitant Variables," *Biometrika*, **63, 3**, 631-8.

Casella, C. and Berger, R. (1990), *Statistical Inference.* California: Wadsworth, Inc.

Cheung, S. H. and Holland, B. (1991), "Extension of Dunnett's Multiple Comparison Procedure to the Case of Several Groups," *Biometrics*, **47**, 21-32.

Cheung, S. H. and Holland, B. (1992), "Extension of Dunnett's Multiple Comparison Procedure with Differing Sample Sizes to the Case of Several Groups," *Computational Statistics and Data Analysis*, **14**, 165-182.

Cheung, S. H. and Holland, B. (1994), "A Step-down Procedure for Multiple Tests of Treatment Versus Control in Each of Several Groups," *Statistics in Medicine*, **13**, 2261-2267.

Cheung, S. H. and Chan, W. S. (1996), "Simultaneous Confidence Intervals for Pairwise Multiple Comparisons in a Two-Way Unbalanced Design," *Biometrics*, **52**, 463-472.

Copenhaver, M. D. (1987), "Multiple Comparisons of Simple Effects in a Two-Way Analysis of Variance with Fixed Effects," Unpublished Ph. D. dissertation, Department of Statistics, Temple University, Philadelphia, Pennsylvania.

Copenhaver, M. D. and Holland, B. (1987), "Calculation of the True Experimentwise Error Rates for Multiple Comparisons Procedures in a Two-Way Factorial Design," *Proceedings, 19th Symposium in the Interface of Computer Science and Statistics*, Philadelphia, Pennsylvania.

Copenhaver, M. D. and Holland, B. (1988), "Computation of the Distribution of the Maximum Studentized Range Statistic with Application to Multiple Significance Testing of Simple Effects," *J. Statist. Comput. Simul.*, **30**, 1-15.

Cox, D. R. (1965), "A Remark on Multiple Comparison Methods," *Technometrics*, **2**, 149-156.

Diaconis, P. (1985), "Theories of Data Analysis from Magical Thinking Through Classical Statistics," in *Exploring Data Tables, Trends and Shapes* (Eds. D. C. Hoaglin, F. Mosteller, and J. W. Tukey). New York: Wiley.

Dunnett, C. W. (1955), "A Multiple Comparison Procedure for Comparing Several Treatments with a Control." *Journal of American Statistical Association*, **50**, 1096-1121.

Fisher, R. A. (1935), *The Design of Experiments.* London: Oliver & Boyd.

Genz, A. (1992), "Numerical Computation of Multivariate Normal Probabilities." *Journal of Computational & Graphical Statistics, Vol. 1,* **No. 2**, 141-149.

Hayter, A. J. (1989), "Pairwise Comparisons of Generally Correlated Means," *Journal of American Statistical Association*, **84**, 208-213.

Hochberg, Y. and Tamhane, A. C. (1987), *Multiple Comparison Procedures.* New York: Wiley.

Hochberg, Y. and Varon-Salomon, Y. (1984), "On Simultaneous Pairwise Comparisons in Analysis of Covariance," *Journal of American Statistical Association*, **79**, 863-866.

Huitema, B. E. (1980), *The Analysis of Covariance and Alternatives.* New York: Wiley.

Joe, H. (1995), "Approximations to Multivariate Normal Rectangle Probabilities Based on Conditional Expectations," *Journal of American Statistical Association*, **90**, 957-964.

Kramer, C. Y. (1956), "Extension of Multiple Range Test to Group Means with Unequal Numbers of Replications," *Biometrics*, **12**, 307-310.

Kramer, C. Y. (1957), "Extension of Multiple Range Test to Group Correlated Adjusted Means," *Biometrics*, **13**, 13-18.

Miller, R. G. (1981), *Simultaneous Statistical Inference*, 2nd Edition. New York: Springer-Verlag.

Neter, J., Wasserman, W. and Kutner, M. H. (1990), *Applied Linear Statistical Models*, 3rd Edition. Illinois: Richard D. Irwin, Homewood.

Scheffé, H. (1953), "A Method for Judging All Contrasts in the Analysis of Variance," *Biometrika*, **40**, 87-104.

Scheffé, H. (1959), *The Analysis of Variance.* New York: Wiley.

Schervish, M. (1984), "Multivariate Normal Probabilities with Error Bound," *Applied Statistics*, **33**, 81-87.

Searle, S. R. (1971), *Linear Models.* New York: Wiley.

Steel, R. G. D. and Torrie, J. H. (1980), *Principles and Procedures of Statistics: A Biometrical Approach,* 2nd Edition. New York: McGraw-Hill.

Thigpen, C. C. and Paulson, A. S. (1974), "A Multiple Range Test for Analysis of Covariance," *Biometrika*, **61, 3**, 479-484.

Tukey, J.W. (1953), "The Problem of Multiple Comparisons," Unpublished manuscript.

Tukey, J.W. (1977), "Some Thoughts on Clinical Trials, Especially Problems of Multiplicity," *Science*, **198**, 679-684.

Westfall, P.H. and Young, S.S. (1993), *Resampling-based Multiple Testing : Examples and Methods for P-value Adjustment.* New York: Wiley.